ÚLOHA O SETKÁNÍ A JEJÍ ZOBECNĚNÍ

LUCIE KREUZIGEROVÁ A ONDŘEJ VENCÁLEK

ABSTRAKT. Článek pojednává o klasické úloze o setkání dvou přátel, k jejímuž řešení se používá tzv. geometrické pravděpodobnosti. Úlohu zobecňujeme a řešíme pro $k \geq 2$ přátel. Kromě exaktního řešení je ukázáno i řešení simulační, dostupné i lidem bez širšího matematického vzdělání.

1. ÚVOD

S *úlohou o setkání* se dnes setká prakticky každý student základního vysokoškolského kurzu pravděpodobnosti obsahujícího pojednání o tzv. geometrické pravděpodobnosti. Ať vezmeme skripta používaná v Brně [1], v Liberci [2], v Olomouci [3], v Ostravě [4], v Praze [5] či v jiném městě, úlohu o setkání tam v kapitole o geometrické pravděpodobnosti s velkou pravděpodobností najdeme. Někdy jde o setkání dvou přátel, jindy to může být setkání dvou vlaků či lodí. Podstata úlohy je však stále stejná. My zde úlohu o setkání formulujeme podobně, jako to činí Anděl v monografii [6]:

Dva přátelé (říkejme jim pan X a pan Y) se domluvili, že přijdou na určité místo v době mezi polednem a jednou hodinou odpoledne. Na místo přijde v tomto časovém intervalu každý z nich zcela náhodně a nezávisle na příchodu toho druhého. Oba jsou připraveni čekat 10 minut, ale ne déle než do jedné hodiny odpoledne. Jaká je pravděpodobnost, že se za těchto podmínek skutečně sejdou?

V tomto článku se zamyslíme nad situací, kdy se mají (či nemají) potkat více než dva přátelé. Představme si, že se na schůzce dohodli tři přátelé. Jaká je nyní pravděpodobnost, že se všichni setkají? A jaká je pravděpodobnost, že se setkají aspoň dva z nich? A jaká je naopak pravděpodobnost, že se nesetkají vůbec? Lze si přitom představit i situace, kdy je žádoucí, aby k žádnému setkání nedošlo. Jde například o situaci, kdy se tři spolubydlící probouzejí nezávisle na sobě mezi 6:00 a 7:00 ráno a poté se odeberou do koupelny, kde se každý z nich zdrží 10 minut. Zajímá nás, jaká je pravděpodobnost, že ani jeden z nich nebude muset čekat. Řešíme tedy spíše „úlohu o nesetkání“.

Úlohu o (ne)setkání pro více než dva lidi řešíme v článku nejprve exaktně, poté své výsledky srovnáme se simulační studií, kterou je schopen realizovat i student, který neumí integrovat (pokud ovšem umí simulovat).

2010 *MSC.* Primární 60A05, 60D05.
Klíčová slova. geometrická pravděpodobnost, setkání, simulace.
Práce byla podporována projektem OP Vzdělávání pro konkurenceschopnost (CZ.1.07/2.3.00/20.0170).

2. Řešení úlohy o setkání

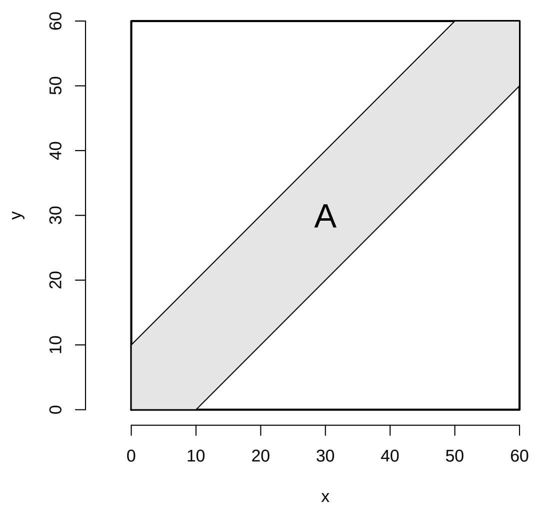
Věnujme se nejprve řešení standardní úlohy o setkání, k němuž se využívá tzv. geometrická pravděpodobnost. Symbolem x označme čas příchodu pána X vyjádřený v nějakých časových jednotkách, například v minutách, uplynulých od poledne. Je třeba si uvědomit, že čas je spojitá veličina, bez ohledu na to, v jakých časových jednotkách je měřen. Proto x může být libovolné reálné číslo z intervalu $[0, 60]$. Žádná hodnota přitom není preferována (není více pravděpodobná než jiné hodnoty). Elementární jev, že pán X přijde v čase x a pán Y v čase y budeme značit (x, y). Množinu všech elementárních jevů, tradičně označovanou symbolem Ω, tedy můžeme zapsat následovně:

$$\Omega = \{(x, y) : x \in [0, 60], y \in [0, 60]\} = [0, 60] \times [0, 60].$$

Jev, jehož pravděpodobnost chceme určit, tedy jev, že se pánové X a Y setkají, zapíšeme následovně:

$$A = \{(x, y) \in \Omega : |x - y| < 10\}.$$

Graficky je celá situace znázorněna na Obrázku 1.



Obrázek 1. Schéma řešení úlohy o setkání. Doba příchodu pána X na x-ové ose, doba příchodu pána Y na y-ové ose. Jev setkání A vyznačen šedě..

Pravděpodobnost jevu A můžeme spočítat jako podíl

$$P(A) = \frac{\lambda(A)}{\lambda(\Omega)},$$

kde symbol λ značí (v tomto případě dvourozměrnou) Lebesgueovu míru. Pravděpodobnost jevu A tedy vypočteme jako podíl šedé plochy při diagonále ku ploše celého čtverce:

$$P(A) = \frac{60^2 - 50^2}{60^2} = \frac{11}{36} \doteq 0,306.$$

Pravděpodobnost, že se dva zmínění pánové setkají, je tedy přibližně 30 %. Naopak k „nesetkání“ dojde s pravděpodobností téměř 70 %.

3. Do hry vstupuje pán Z

Představme si nyní, že se s pány X a Y chce setkat ještě pán Z. Zde může nastat vícero situací: buďto se sejdou všichni, nebo se sejdou jen dva z nich a třetího přitom minou, nebo se nesetkají vůbec, tedy žádný z pánů nepotká žádného ze svých přátel. Položme si otázku, jaká je pravděpodobnost, že dojde právě k této situaci, tedy že dojde k „nesetkání“.

Princip řešení úlohy je stejný. Prostorem elementárních jevů je nyní množina

$$\Omega = [0, 60]^3,$$

jeho prvky značíme (x, y, z). Aspoň dva ze tří pánů se setkají, pokud platí aspoň jedna z následujících nerovností: $|x - y| < 10$, $|x - z| < 10$, $|y - z| < 10$. V opačném případě dojde k nesetkání. Tento jev tedy můžeme popsat následovně:

$$N = \{(x, y, z) \in \Omega : \min(|x - y|, |x - z|, |y - z|) > 10\}.$$

K určení pravděpodobnosti nesetkání tedy stačí vypočíst

$$P(N) = \frac{\lambda(N)}{\lambda(\Omega)}.$$

Nakreslit obrázek popisující tuto situaci již bude trochu komplikovanější (obrázek je trojrozměrný) a výpočet vyžaduje buďto velmi dobrou geometrickou představivost nebo schopnost integrovat.

Pro větší obecnost budeme uvažovat, že délka časového úseku, v němž se přátelé mají (nemají) setkat je d (v našem případě $d = 60$ minut) a délka, po kterou jsou ochotni čekat je c (v našem případě $c = 10$ minut). Budeme předpokládat, že $d > 2c$, neboť v opačném případě by k nesetkání tří osob vůbec nemohlo dojít.

Situaci si značně zjednodušíme, pokud budeme předpokládat nějaké uspořádání hodnot x, y a z, například $x < y < z$. Toto uspořádání je stejně pravděpodobné, jako kterékoliv jiné uspořádání, tedy jeho pravděpodobnost je $1/(3!)$. Lebesgueovu míru části množiny N, pro níž je splněna podmínka $x < y < z$ pak vypočteme jako integrál

$$\int_0^{d-2c} \int_{x+c}^{d-c} \int_{y+c}^{d} \mathrm{d}z\, \mathrm{d}y\, \mathrm{d}x = (d - 2c)^3/6,$$

tedy Lebesgueova míra celé množiny N bude rovna $(d - 2c)^3$ a pravděpodobnost jevu N můžeme vypočíst jako podíl

$$P(N) = \frac{(d - 2c)^3}{d^3} = \left(1 - 2\frac{c}{d}\right)^3.$$

Povšimněme si, že výsledná pravděpodobnost záleží jen na poměru c/d, tedy na tom, jakou část z celkové doby určené na setkání jsou jednotliví lidé ochotni čekat.

Při větším počtu osob, které se mají (nemají) setkat, lze postupovat obdobně. Pro $2 \le k \le d/c$ osob platí, že pravděpodobnost jejich nesetkání je rovna

$$P(N) = \left(1 - (k-1)\frac{c}{d}\right)^k .$$

Vrátíme-li se k situaci, kdy $d = 60$ a $c = 10$, můžeme postupně uvažovat o 2 až 6 osobách. Pravděpodobnosti jejich nesetkání jsou shrnuty v tabulce 1.

k	2	3	4	5	6
P(N)	25/36	8/27	1/16	1/243	1/46656

Tabulka 1. Pravděpodobnosti nesetkání pro různé počty osob.

Zatímco při poměru $c/d = 1/6$ se dvě osoby spíše nepotkají (s pravděp. 70 %), u tří osob už nesetkání nastává s pravděpodobností jen asi 30 %, u čtyř s pravděpodobností asi 6 % a u více než čtyř už je pravděpodobnost velmi malá. Pro více než 6 osob už je pravděpodobnost nesetkání nulová.

4. Setkání a nesetkání

V předešlé sekci jsme se zabývali „nesetkáním", tedy situací, kdy ani jedna z osob, které se mají setkat, nepotká žádnou další osobu. Opakem tohoto jevu je jev, kdy dojde k sektání alespoň dvou osob. Pravděpodobnost tohoto jevu je tedy doplňkem do jedničky z výše uvedené pravděpodobnosti nesetkání.

Mohla by nás však zajímat i pravděpodobnost jevu, že se sejdou všichni přátelé. Zde je dobré si uvědomit, jak ve skutečnosti vypadá situace, kdy se dva z přátel opravdu sejdou. Můžeme uvažovat o různých scénářích v chování těchto osob. Pokud je bodem setkání například místo, kde se mohou občerstvit, budou zřejmě ochotni se společně občerstvovat a čekat na další účastníky schůzky delší dobu než jen stanovených 10 minut (třeba až do konce dohodnuté či otevírací doby).

Představme si však, že každý z účastníků má na setkání vyhrazeno právě jen 10 minut. Jaká je pravděpodobnost, že se setkají všichni? Jev setkání všech účastníků můžeme pro tři osoby zapsat následovně:

$$S = \{(x,y,z) \in \Omega : \max(|x-y|, |x-z|, |y-z|) < 10\} .$$

Pravděpodobnost tohoto jevu opět vypočteme pomocí podílu

$$P(S) = \frac{\lambda(S)}{\lambda(\Omega)}.$$

Znovu si pomůžeme předpokladem $x < y < z$. Lebesgueovu míru části množiny S, pro níž je splněna podmínka $x < y < z$, pak vypočteme jako součet integrálů

$$\int_0^{d-c} \int_x^{x+c} \int_y^{x+c} \mathrm{d}z\, \mathrm{d}y\, \mathrm{d}x + \int_{d-c}^{d} \int_x^{d} \int_y^{d} \mathrm{d}z\, \mathrm{d}y\, \mathrm{d}x = c^2(3d-2c)/6.$$

Pravděpodobnost jevu S tedy vypočteme následovně:

$$P(S) = \frac{c^2(3d-2c)}{d^3} = 3\,(c/d)^2 - 2\,(c/d)^3 .$$

Výsledná pravděpodobnost opět záleží jen na poměru c/d, tedy na tom, jakou část z celkové doby určené na setkání jsou jednotliví lidé ochotni čekat.

Při větším počtu osob, které se mají setkat, lze postupovat obdobně. Za zmínku snad stojí skutečnost, že nyní může k setkání všech osob dojít pro libovolně velký počet osob (bez omezení), i když pochopitelně pravděpodobnost bude s rostoucím počtem účastníků setkání klesat. Pro $k \geq 2$ osob platí, že pravděpodobnost setkání všech je rovna

$$P(S) = k\,(c/d)^{k-1} - (k-1)\,(c/d)^{k}\,.$$

V situaci, kdy $d = 60$ a $c = 10$, jsou pravděpodobnosti setkání všech osob shrnuty v tabulce 2.

k	2	3	4	5	...	10	...
P(S)	11/36	2/27	7/432	13/3 888	...	17/20 155 392	...

Tabulka 2. Pravděpodobnosti setkání pro různé počty osob.

Při poměru $c/d = 1/6$ se dvě osoby setkají s pravděpodobností přibližně 30 % (11/36), tři osoby s pravděpodobností jen asi 7 % (2/27). Čím více je osob, které se mají setkat, tím je pravděpodobnost, že se potkají všechny, menší. Je však nenulová pro libovolně velký (konečný) počet osob.

5. Simulace nesetkání

Úloha o (ne)setkání je atraktivní také v tom, že se dá velmi snadno řešit simulačně. Pomocí generátoru (pseudo)náhodných čísel opakovaně generujeme z rovnoměrného rozdělení časy příchodů jednotlivých pánů (nezávisle na sobě) a poté vzhodnocujeme, zda došlo k (ne)setkání. Relativní četnost situací, kdy došlo k „nesektání“, je pak odhadem pravděpodobnosti tohoto jevu.

Ukažme si kód takovéto simulace realizovaný v prostředí statistického softwaru R. Nejprve se věnujme situaci, kdy se (ne)mají setkat jen dvě osoby:

```
opak = 100000  # pocet opakovani simulace
d    = 60      # delka casoveho useku
c    = 10      # doba cekani

x1 = runif(opak,0,d)  # casy prichodu pana X
x2 = runif(opak,0,d)  # casy prichodu pana Y

pravdep = sum(abs(x1-x2)>c)/opak
print(pravdep)
```

Naše simulace při 100 000 opakováních (výsledek máme prakticky ihned) dospěla k odhadu pravděpodobnosti nesetkání 0,69414, což je dosti blízko skutečné hodnotě 0,69444. Svou nejistotu v odhadu přitom můžeme snadno vyjádřit třeba pomocí 95%-ho (přibližného) konfidenčního intervalu. K jeho sestrojení použijeme centrální limitní větu. V našem případě má tento konfidenční interval podobu $(0,691284; 0,696996)$. Vidíme tedy, že již třetím desetinným místem si nemůžeme být moc jisti. Zpřesnění o řád si přitom vyžádá stokrát více opakování simulace. Zde je také vidět, že simulací budeme jen velmi těžko odhadovat nízké pravděpodobnosti, třeba pravděpodobnost nesetkání pro 5 nebo 6 osob (když $c/d = 1/6$).

Věnujme se nyní úloze (ne)setkání více osob, přičemž chceme simulaci naprogramovat tak, aby fungovala pro libovolný počet $k \leq d/c$ lidí, kteří se mají (nemají) setkat. Je třeba si uvědomit, že k vyhodnocení, zda došlo k nesetkání, obecně potřebujeme porovnat všechny dvojice časů příchodů. Těch je $k(k-1)/2$. Můžeme sice nejprve časy uspořádat a poté provést jen $k-1$ porovnání, uvědomme si však, že opakované seřazení hodnot od nejmenší k největší už je poměrně náročné na čas.

Přímočará simulace má sice přehledný zápis, ale její výpočet je poněkud zdlouhavý:

```
opak = 100000  # pocet opakovani simulace
d    = 60      # delka casoveho useku
c    = 10      # doba cekani
k    = 4       # pocet osob

nesetkajise = 0
for (i in 1:opak) {
  prichody = runif(k,0,d)
  prichody = sort(prichody)
  rozdily  = diff(prichody)
  if (min(rozdily) > c) nesetkajise = nesetkajise + 1 }

pravdep = nesetkajise/opak
print(pravdep)
```

Tato simulace odhadne pravděpodobnost nesetkání v případě 4 osob, které se mají setkat během jedné hodiny, když každá z osob je ochotna čekat 10 minut. Na počítači s procesorem 2.66GHz Intel Core 2 trvá při 100 000 opakováních asi 13 sekund. Oproti tomu následující verze využívá dobrých vlastností softwaru R při práci s vektory. Při stejných parametrech trvá jen asi 1 sekundu. Úloha o nesetkání tedy nemusí být napříště spojována jen s geometrickou pravděpodobností, ale stejně dobře může posloužit i k zdokonalování programátorských schopností studentů.

```
# matice pro ulozeni casu prichodu:
x = matrix(ncol=k,nrow=opak)

# generovani casu prichodu:
for (i in 1:k)   x[,i] = runif(opak,0,d)

# porovnani casu prichodu:
rozdil = matrix(ncol=choose(k,2),nrow=opak)
index = 1
for (i in 1:(k-1))
  for (j in (i+1):k)
    {rozdil[,index] = abs(x[,i]-x[,j])
```

```
    index = index+1   }

m = apply(rozdil,1,min)
pravdep = sum(m>c)/opak
print(pravdep)
```

6. Závěr

Úloha o sektání dvou přátel se v různých obměnách objevuje snad až příliš často jako ukázkový příklad využití tzv. geometrické pravděpodobnosti. V článku jsme ukázali možnost, jak úlohu zobecnit pro setkání vícero osob. Nalezli jsme exaktní tvar řešení této úlohy a poukázali na možnost využití simulace k nalezení přibližného řešení. Úloha je zajímavá i z pohledu programátorského.

Reference

[1] M. Budíková, Š. Mikoláš, P. Osecký: *Teorie Pravděpodobnosti a matematická statistika – sbírka příkladů*, 3. vydání, Masarykova univerzita, Brno, 2007.
[2] L. Moc: *Cvičení z inženýrské statistiky (řešené příklady)*, Technická univerzita v Liberci, 2006.
[3] K. Hron, P. Kunderová: *Základy počtu pravděpodobnosti a metod matematické statistiky*, Univerzita Palackého, Olomouc, 2013.
[4] P. Otipka, V. Šmajstrla: *Pravděpodobnost a statistika [elektronický zdroj]*, Technická univerzita Ostrava, 2008.
[5] V. Dupač, M. Hušková: *Pravděpodobnost a matematická statistika*, Karolinum, Praha, 2003.
[6] J. Anděl: *Matematika náhody*, Matfyzpress, Praha, 2003.

Lucie Kreuzigerová, Katedra matematické analýzy a aplikací matematiky, Přírodovědecká fakulta, Univerzita Palackého v Olomouci, 17. listopadu 12, 771 46 Olomouc, Česká republika, *e-mail*: `lucie.kreuzigerova01@upol.cz`

Ondřej Vencálek, Katedra matematické analýzy a aplikací matematiky, Přírodovědecká fakulta, Univerzita Palackého v Olomouci, 17. listopadu 12, 771 46 Olomouc, Česká republika, *e-mail*: `ondrej.vencalek@upol.cz`